On tail trend detection: modeling relative risk*

Laurens de Haan
University of Tilburg
University of Lisbon
Erasmus University Rotterdam

Albert Klein Tank
KNMI, Royal Netherlands
Meteorological Institute

Cláudia Neves
CEAUL and University of Aveiro

Abstract

A simple and natural model is introduced for studying a trend in the tail of a probability distribution over time (or over space). Estimation and testing procedures are provided with application to rainfall data.

KEY WORDS AND PHRASES: extreme value distribution, regular variation, extreme rainfall

1 Introduction

In the climate change dispute some people suggest (Klein Tank and Können (2003); Groisman et al. (2005); Alexander et al. (2006); Zolina et al. (2009)) that perhaps there is no or little change in the mean of the probability distribution of daily rainfall over time but there is a change in the tail that is, more extreme events occur more frequently. The present paper – like Smith (1989); Hall and Tajvidi (2000); Hanel et al. (2009) – considers a trend in extremes from the point of view of extreme value theory. A simple and intuitive way to study such a trend is by looking at the tail of the probability distribution at time $s > 0$ compared to the tail at time zero.

We consider random variables $X(s)$ where $s \geq 0$ is time. Write $F_s(x) := P\{X(s) \leq x\}$ for $x \in \mathbb{R}$. Assume that F_0 is in the domain of attraction of an extreme value distribution ($F_0 \in \mathcal{D}(G_\gamma)$ for some $\gamma \in \mathbb{R}$) and that

$$\frac{1 - F_s(x)}{1 - F_0(x)}$$

tends to a positive constant for all $s > 0$ when x tends to the right endpoint x^* of F_0. Hence the exceedance probability at time s is systematically a factor times the exceedance probability at time 0. We take for simplicity

*Research partially supported by ENES–Project PTDC/MAT/112770/2009.

as our model for relative risk

$$\lim_{x\uparrow x^*}\frac{1-F_s(x)}{1-F_0(x)}=e^{cs}, \tag{1}$$

with c a real constant representing a possible trend. This means that e.g. for $s=1$ and $c=1$ the probability of an extreme event taking place at time 1 is e times the probability at time 0.

First note that by Resnick (1987), section 1.5, relation (1) implies that $F_s \in \mathcal{D}(G_\gamma)$ for all $s \geq 0$. Recall that $F_s \in \mathcal{D}(G_\gamma)$ if and only if for some positive function a_s

$$\lim_{t\to\infty}\frac{U_s(tx)-U_s(t)}{a_s(t)}=\frac{x^\gamma-1}{\gamma}, \tag{2}$$

$x>0$, where U_s is the inverse function of $1/(1-F_s)$.

A slight extension of Resnick's result provides a relation equivalent to (1) for the functions U_s: for $s>0$

$$\lim_{t\to\infty}\frac{U_s(t)-U_0(t)}{a_0(t)}=\frac{e^{c\gamma s}-1}{\gamma}. \tag{3}$$

Note that relation (3) implies

$$\lim_{t\to\infty}\frac{a_s(t)}{a_0(t)}=e^{c\gamma s}.$$

Furthermore, for $\gamma>0$, relation (3) is equivalent to

$$\lim_{t\to\infty}\frac{U_s(t)}{U_0(t)}=e^{c\gamma s}. \tag{4}$$

Relations (1), (3) and (4) can be used to build estimators for c. We proceed in a semi-parametric way.

Suppose that we have repeated observations at discrete time points $0=s_0<s_1<s_2<\ldots<s_m$. It is assumed that $\left\{X_i(s_j)\right\}_{i=1}^{n}{}_{j=1}^{m}$ are all independent and that $X_1(s_j)$, $X_2(s_j),\ldots,X_n(s_j)$ have the same distribution function F_{s_j} for all j. Let $X_{1,n}(s_j)\leq X_{2,n}(s_j)\leq\ldots\leq X_{n,n}(s_j)$ be their order statistics.

(i) If it is known that γ is positive, relation (4) leads to a least square estimator for c:

$$\hat{c}^{(1)}:=\frac{\sum_{j=1}^{m}s_j\left(\log X_{n-k,n}(s_j)-\log X_{n-k,n}(0)\right)}{\hat{\gamma}^+_{n,k}\sum_{j=1}^{m}s_j^2}. \tag{5}$$

We shall discuss the estimator $\hat{\gamma}^+_{n,k}$ for $\gamma^+:=\max(0,\gamma)$ later on.

(ii) Relation (3) also leads to an estimator for c. Intuitively relation (3) means that

$$\left(1+\hat{\gamma}_{n,k}\frac{X_{n-k,n}(s_j)-X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}}\approx e^{cs_j},$$

where $\hat{\gamma}_{n,k}$ is an estimator for γ. Define $\hat{c}^{(2)}$ by

$$\arg\min_c \sum_{j=1}^{m} \left\{ \log\left(1 + \hat{\gamma}_{n,k} \frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}} - cs_j \right\}^2$$

i.e.,

$$\hat{c}^{(2)} := \frac{\sum_{j=1}^{m} s_j \ \log\left(1 + \hat{\gamma}_{n,k} \ \frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0(n/k)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}}}{\sum_{j=1}^{m} s_j^2}. \tag{6}$$

For $\hat{\gamma}_{n,k} = 0$, the estimator is defined by continuity.

(iii) Finally relation (1) leads to an estimator for c. Intuitively relation (1) means that

$$\log \frac{1 - \widehat{F}_s\big(X_{n-k,n}(0)\big)}{1 - \widehat{F}_0\big(X_{n-k,n}(0)\big)} \approx cs$$

where $\widehat{F}_s$ is the empirical distribution function at time s. This leads to

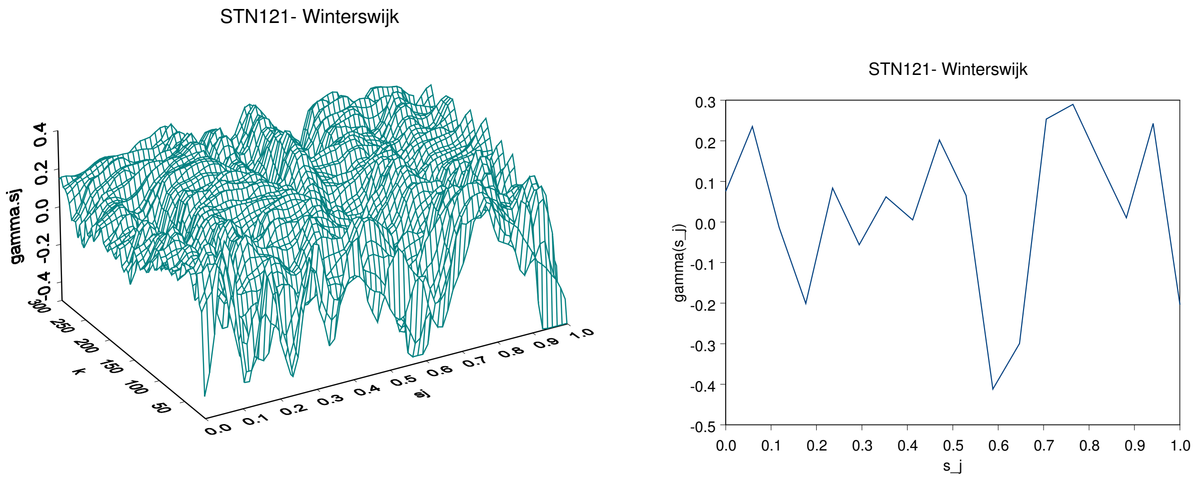
$$\hat{c}^{(3)} := \frac{\sum_{j=1}^{m} \log\left(\frac{1}{k} \sum_{i=1}^{n} I_{\{X_i(s_j) > X_{n-k,n}(0)\}}\right)}{\sum_{j=1}^{m} s_j}. \tag{7}$$

The problem of defining and estimating a trend in extreme value theory has been considered by a number of authors including Smith (1989); Hall and Tajvidi (2000); Coles (2001); Yee and Stephenson (2007) and more recently addressed by Mannshardt-Shamseldin et al. (2010). A review of results on trend estimation is given in the Appendix.

We have restricted ourselves to the model in (1) that is, a trend function of the form e^{cs}, since we are interested in a monotone trend and also because a more general change would have been more difficult to detect. Figure 1 gives some insight into the difficulty of detecting more complex trend functions, namely a temporal trend in the extreme value index γ. We could have studied more general (not monotone) changes in a similar manner, possibly with adjustments enabling other appropriate estimation procedures.

The outline of this paper is as follows. In section 2 we state the conditions for consistency and asymptotic normality of the estimators introduced in (5), (6) and (7). Proofs are postponed to section 5. In section 3 we collect some simulation results for illustrating and assessing finite sample performance of the various estimators for the trend. In section 4 we apply the methods to daily rainfall at 18 stations across Germany and The Netherlands and give a tentative interpretation of the results. Indeed for some stations the probability of extreme rainfall has increased by about 2% in each decade.

Figure 1: Estimates $\hat{\gamma}(s_j)$ for each one of the $m+1 = 18$ time points $s_j = j/m$, $j = 0, 1, 2, \ldots, m$, at one particular location, a gauging station in The Netherlands. The number k corresponds to the number of observations above the hight random threshold $X_{n-k,n}(s_j)$ (Left). Corresponding estimates $\hat{\gamma}(s_j)$ with $k = 30$ (Right).



2 Results

Let us consider $\hat{c}^{(1)}$ first and suppose $\gamma > 0$. Let $\hat{\gamma}^+_{n,k}(s_j)$ be an estimator for γ^+ based on the observations at time s_j. Since γ^+ does not depend on s, we use a combined estimator

$$\hat{\gamma}^+_{n,k} := \frac{1}{m} \sum_{j=1}^{m} \hat{\gamma}^+_{n,k}(s_j). \tag{8}$$

We consider estimators $\hat{\gamma}^+_{n,k}(s_j)$ that are consistent, i.e., if (4) holds and the number $k = k_n$ of upper order statistics used satisfies $k = k_n \to \infty$, $k_n/n \to 0$, $n \to \infty$, then $\hat{\gamma}^+_{n,k}(s_j) \xrightarrow[n\to\infty]{P} \gamma^+$.

Next we discuss the asymptotic normality of $\hat{\gamma}^+_{n,k}$. Consider a second order condition for U_0: suppose there exists a positive or negative function β with $\lim_{t\to\infty} \beta(t) = 0$ such that for $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_0(tx)}{U_0(t)} - x^{\gamma^+}}{\beta(t)} = x^{\gamma^+} \frac{x^{\widetilde{\rho}} - 1}{\widetilde{\rho}} \tag{9}$$

with $\widetilde{\rho}$ a non-positive parameter. Further we need a second order strengthening of condition (4): suppose that for all j

$$\lim_{t\to\infty} \frac{\frac{U_{s_j}(t)}{U_0(t)} - e^{c\gamma^+ s_j}}{\beta(t)} = e^{c\gamma^+ s_j} \frac{e^{c\widetilde{\rho} s_j} - 1}{\widetilde{\rho}}. \tag{10}$$

We consider estimators $\hat{\gamma}^+_{n,k}(s_j)$ such that

$$\sqrt{k}\Big(\hat{\gamma}^+_{n,k}(s_j) - \gamma^+,\ \log X_{n-k,n}(s_j) - \log U_{s_j}\big(\frac{n}{k}\big)\Big) \xrightarrow[n\to\infty]{d} \big(\Gamma^+(s_j),\ B^+(s_j)\big), \tag{11}$$

say, for all j, where $\big(\Gamma^+(s_j),\ B^+(s_j)\big)$ has a multivariate normal distribution provided k (the number of upper order statistics used in $\hat{\gamma}^+_{n,k}(s_j)$ for all j) satisfies $k = k_n \to \infty$ and

$$\lim_{n\to\infty} \sqrt{k_n}\, \beta\Big(\frac{n}{k_n}\Big) =: \lambda$$

exists finite. Various estimators $\hat{\gamma}^+_{n,k}(s_j)$ are known with this property, notably Hill's estimator (Hill (1975)).

Next we consider $\hat{c}^{(2)}$. Consider for each j estimators $\hat{\gamma}_{n,k}(s_j)$, $\hat{a}_{s_j}(n/k)$ with the following properties.

a. Under the conditions $F_0 \in \mathcal{D}(G_\gamma)$ and (1)

$$\hat{\gamma}_{n,k}(s_j) \xrightarrow[n\to\infty]{P} \gamma, \quad \frac{\hat{a}_{s_j}\big(\frac{n}{k}\big)}{a_{s_j}\big(\frac{n}{k}\big)} \xrightarrow[n\to\infty]{P} 1 \tag{12}$$

provided $k = k_n \to \infty$, $k_n/n \to 0$, $n \to \infty$.

b. Consider the following second order strengthening of the two conditions in **a**: there exists a positive or negative function α with $\lim_{t\to\infty} \alpha(t) = 0$ such that for each j and $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_0(tx)-U_0(t)}{a_0(t)} - \frac{x^\gamma - 1}{\gamma}}{\alpha(t)} = \frac{1}{\rho}\bigg(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\bigg) =: H_{\gamma,\rho}(x) \tag{13}$$

where ρ is a non-positive parameter. Further:

$$\lim_{t\to\infty} \frac{\frac{U_{s_j}(t)-U_0(t)}{a_0(t)} - \frac{e^{c\gamma s_j} - 1}{\gamma}}{\alpha(t)} = H_{\gamma,\rho}(e^{cs_j}). \tag{14}$$

Under conditions (13) and (14)

$$\sqrt{k}\bigg(\hat{\gamma}_{n,k}(s_j) - \gamma,\ \frac{\hat{a}_{s_j}\big(\frac{n}{k}\big)}{a_{s_j}\big(\frac{n}{k}\big)} - 1,\ \frac{X_{n-k,n}(s_j) - U_{s_j}(n/k)}{a_{s_j}\big(\frac{n}{k}\big)}\bigg)$$
$$\xrightarrow[n\to\infty]{d} \quad \big(\Gamma(s_j),\ A(s_j),\ B(s_j)\big),$$

say, where $\big(\Gamma(s_j), A(s_j),\ B(s_j)\big)$ are independent random vectors and have a multivariate normal distribution for each j provided $k = k_n \to \infty$, and

$$\lim_{n\to\infty} \sqrt{k_n}\, \alpha\Big(\frac{n}{k_n}\Big) =: \lambda$$

exists finite. Various estimators are known with these properties.

Theorem 1 *1. Assume* $\gamma > 0$*. Under condition* (1)

$$\hat{c}^{(1)} \xrightarrow[n\to\infty]{P} c.$$

Under conditions (9) *and* (10)

$$\sqrt{k}\left(\hat{c}^{(1)} - c\right) \xrightarrow[n\to\infty]{d} \frac{\sum_{j=1}^{m} s_j \left(\frac{A(s_j)-A(0)}{\gamma} + \frac{e^{c\tilde{\rho}s_j}-1}{\tilde{\rho}}\,\lambda\right)}{\sum_{j=1}^{m} s_j^2} - \frac{c}{\gamma}\,\frac{1}{m}\sum_{j=1}^{m}\Gamma(s_j).$$

2. Under condition (1)

$$\hat{c}^{(r)} \xrightarrow[n\to\infty]{P} c \quad \text{for } \; r = 2,\, 3.$$

Under conditions (13) *and* (14)

$$\sqrt{k}\left(\hat{c}^{(2)} - c\right) \xrightarrow[n\to\infty]{d} \sum_{j=1}^{m} s_j \Big\{ \frac{1 - e^{-c\gamma s_j} - c\gamma s_j}{\gamma^2}\,\frac{1}{m}\sum_{i=1}^{m}\Gamma(s_i)$$
$$+ e^{-c\gamma s_j}\Big(e^{c\gamma s_j}B(s_j) - B(0) - \frac{e^{c\gamma s_j}-1}{\gamma}\,A(0) + \lambda\, H_{\gamma,\rho}(e^{cs_j})\Big)\Big\} \Big/ \sum_{j=1}^{m} s_j^2$$

and

$$\sqrt{k}\left(\hat{c}^{(3)} - c\right) \xrightarrow[n\to\infty]{d} \sum_{j=1}^{m} \Big\{ e^{-cs_j} W^{(s_j)}(e^{cs_j}) - W^{(0)}(1) + e^{-cs_j} b_3(s_j)\,\lambda \Big\} \Big/ \sum_{j=1}^{m} s_j^2,$$

where $\{W^{(s_j)}(t)\}_{t\geq 0}$ *are independent standard Brownian motions* $(j = 1, 2, \ldots, m)$ *and*

$$b_3(s_j) = \begin{cases} \left(1 - \frac{1}{\rho}\right)\frac{e^{-c\gamma s_j}-1}{\gamma}, & \rho < 0, \\ 0, & \rho = 0. \end{cases}$$

Corollary 2 *Assume* $c = 0$*. Under the conditions of the Theorem,*

1. *if* $k = k_n$ *is such that* $\sqrt{k}\,\beta(n/k) \to 0$*, as* $n \to \infty$ *then*

$$Q^{(1)}_{m,n} := \sum_{j=1}^{m} \frac{k}{2} \left\{ \frac{\log X_{n-k,n}(s_j) - \log X_{n-k,n}(0)}{\hat{\gamma}^{+}_{n,k}} \right\}^2 \xrightarrow[n\to\infty]{d} \chi^2(m); \tag{15}$$

2. *if* $k = k_n$ *is such that* $\sqrt{k}\,\alpha(n/k) \to 0$*, as* $n \to \infty$ *then*

$$Q^{(2)}_{m,n} := \sum_{j=1}^{m} \frac{k}{2} \left\{ \frac{1}{k} \sum_{i=1}^{n} I_{\{X_i(s_j) > X_{n-k,n}(0)\}} - 1 \right\}^2 \xrightarrow[n\to\infty]{d} \chi^2(m). \tag{16}$$

Here $\chi^2(m)$ is a standard chi-squared distributed random variable with m degrees of freedom.

Corollary 2 gives rise to a testing procedure for detecting the presence of a trend in the tail of the underlying distribution functions F_s all lying in the same domain of attraction. That is, $Q^{(r)}_{m,n}$, $r = 1, 2$, defined above can be used as test statistics to evaluate the null hypothesis $H_0 : c = 0$ against the alternative $H_1 : c \neq 0$. Whence H_0 should be rejected at a significance level $\alpha \in (0, 1)$ for any observed value of $Q^{(r)}_{m,n}$ verifying $Q^{(r)}_{obs} > q_{1-\alpha}(m)$, the latter being the $(1 - \alpha)$-quantile pertaining to the chi-squared distribution with m degrees of freedom.

3 Simulations

Simulations have been carried out for three distributions, the generalized Pareto distribution, the ordinary Pareto distribution and the Cauchy distribution. The number of locations is 200 (i.e. $m = 200$) with $s_j = j/m$, $j = 1, 2, \ldots, m$. At each location there are 500 i.i.d. observations ($n = 500$). Then there are 1000 replications.

The generalized Pareto distribution (GPD) with distribution function $1 - (1 + \gamma x)^{-1/\gamma}$ for those x for which $1 + \gamma x > 0$ has been considered with $\gamma = -0.1, 0.1$ and 0.5 (since 0.1 is a typical value for rainfall, the application topic). Relation (13) holds with limit zero since the left hand-side is zero (exact fit). The values $c = -0.1$ and $c = 0.1$ have been considered.

The starting point in each case is a r.v. X from the GPD distribution. For location s_j we then take $X(s_j) \stackrel{d}{=} e^{cs_j\gamma} X + \left(e^{cs_j\gamma} - 1\right)/\gamma$. That way the relations (1) and (14) hold.

Figure 2 displays the average (over 1000 replications) values of the estimators $\hat{c}^{(1)}$ (only for positive γ), $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ as functions of the number k of upper order statistics used in the estimation. As usual in graphs of this type there is a stretch of the graph that is more or less straight; the idea is that in that part both the variance and bias are not too high. The estimators $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ seem to give the best performance.

In Figure 2 the extreme value index γ and scale a_0 have been estimated by the moment estimator. Figure 3 gives a comparison with the maximum likelihood estimator.

For the ordinary Pareto distribution with distribution function $1 - x^{-1/\gamma}$, $x \geq 1$, $\gamma > 0$, we simulate the trend by taking $X(s_j) \stackrel{d}{=} e^{cs_j\gamma} X$ where X follows the Pareto distribution. Again we have an exact fit.

For the Cauchy distribution again the trend is simulated by taking $X(s_j) \stackrel{d}{=} e^{cs_j} X$ (since γ is 1). Relations (13) and (14) hold. In this case $|\alpha|$ is a regularly function with index -2 (rather fast convergence). Figure 4 displays the simulation results for the Pareto and Cauchy distributions.

A much more comprehensive simulation study has been performed. From the described simulations and the other ones we conclude that the estimators perform reasonably well. Estimators $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ seem to behave better than $\hat{c}^{(1)}$. In the next (application) section we shall emphasize the smoother $\hat{c}^{(2)}$.

Figure 2: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$, with underlying Generalized Pareto distribution, in either case of true value $c = -0.1$ or $c = 0.1$.

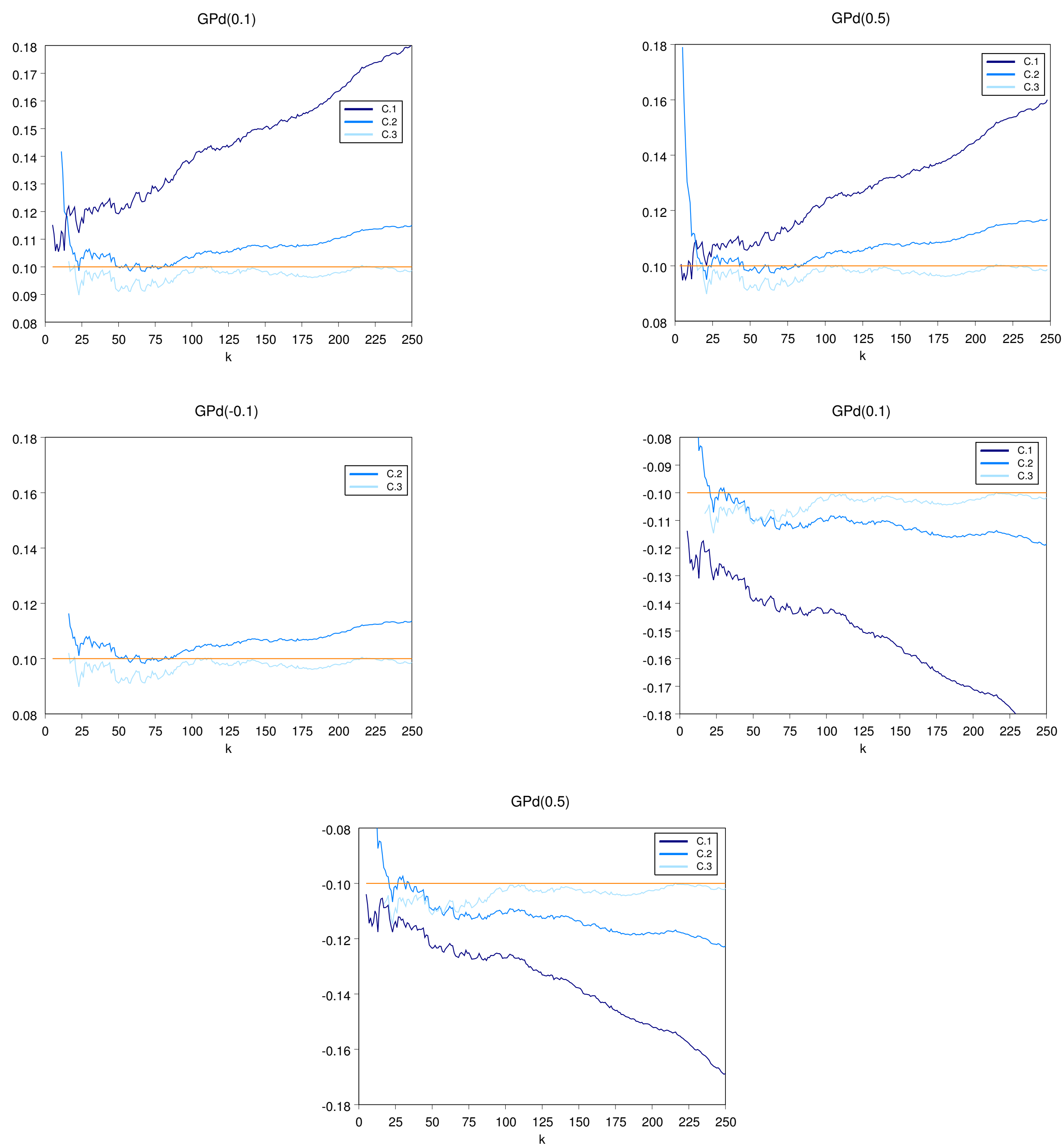
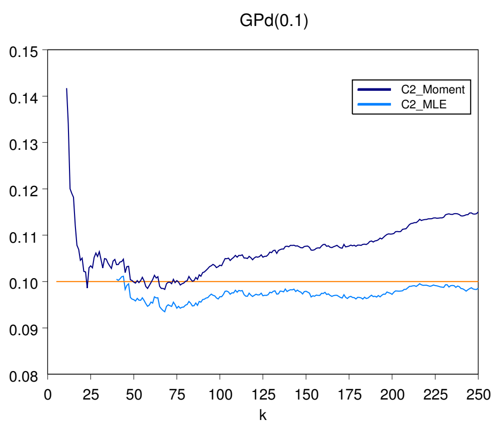


Figure 3: Estimated means of $\hat{c}^{(2)}$ either with Moment estimator and ML estimator for the Generalized Pareto distribution with $\gamma = 0.1$ and $c = 0.1$. The number of samples amongst the 1000 replicates that have produced valid ML-estimates is presented on the right hand-side.



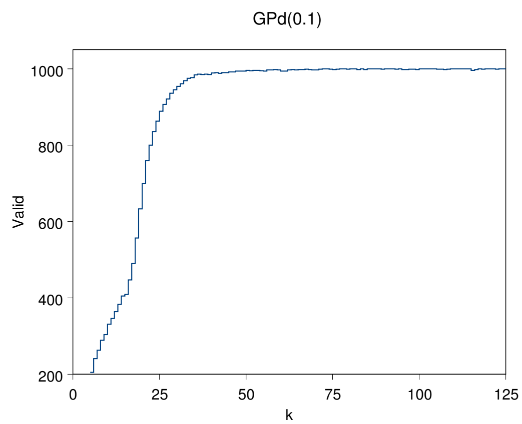


Figure 4: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \dots, m$, in case of true value $c = 0.1$, with Pareto and Cauchy parent distributions.

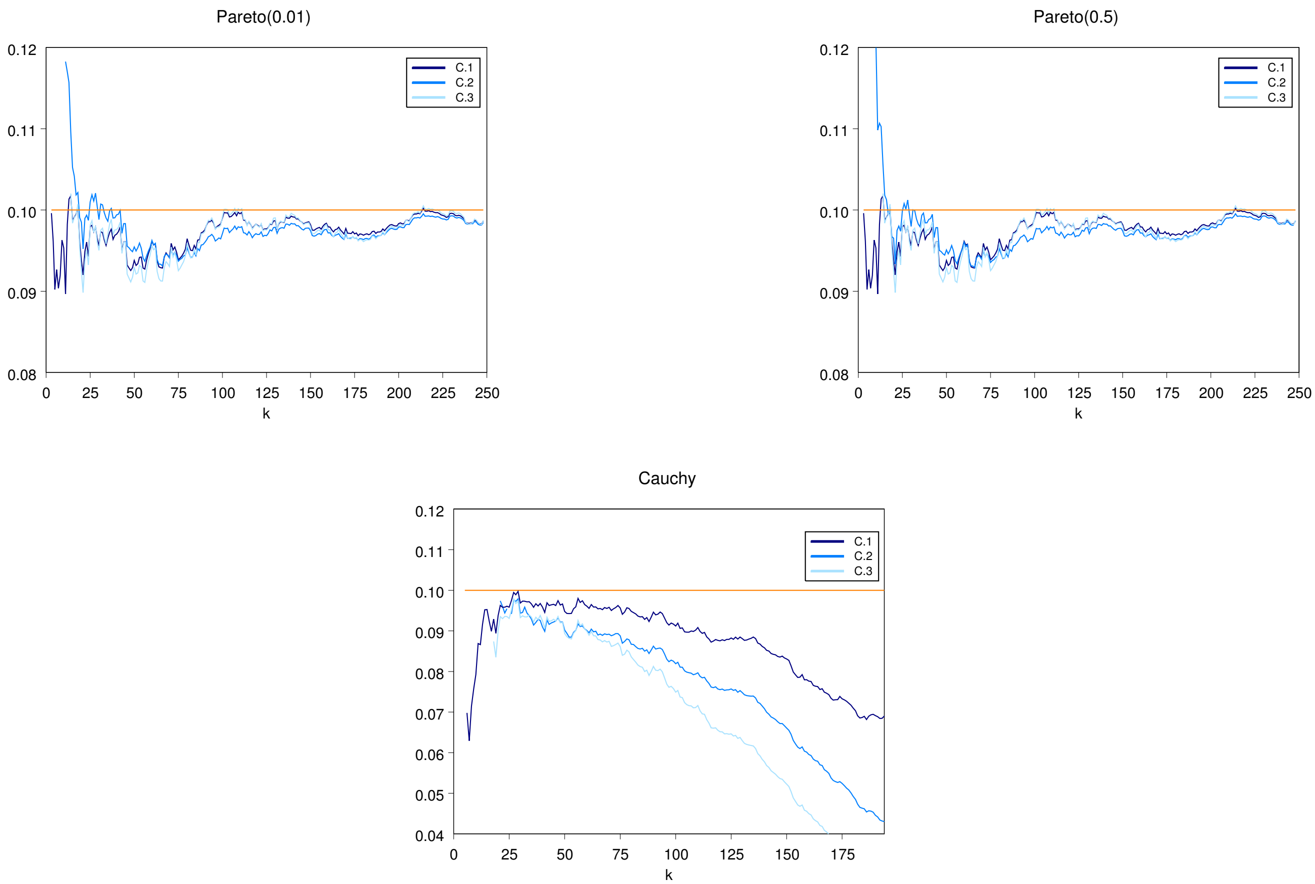
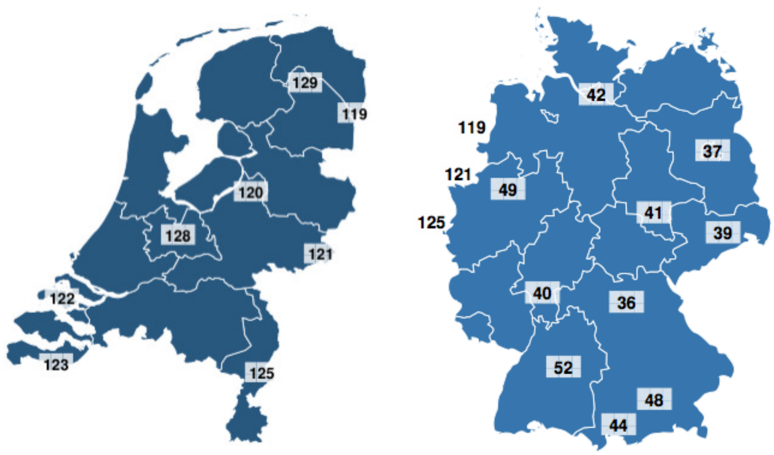


Figure 5: Selected gauging stations in The Netherlands (*left*); selected gauging stations in Germany and Dutch stations near the borderline (*right*).



4 Data Analysis

As an application of the tail trend assessment methodology developed in this paper, we will look at daily rainfall totals collected in 18 gauging stations across Germany and The Netherlands, comprising latitude 47N-53N and longitude 5E-13E. The geographic location of the stations is displayed in Figure 5. Rainfall data are from the European Climate Assessment and Dataset. We note that different stations suffer from different coverage in time in the sense that not all stations have started regular recording of data at the same year. Moreover there are some stations with missing values. All of them however meet the basic criterium for completeness that there is less than 10 days missing per year which leaves us with 90 years of complete data from 1918 up to 2007.

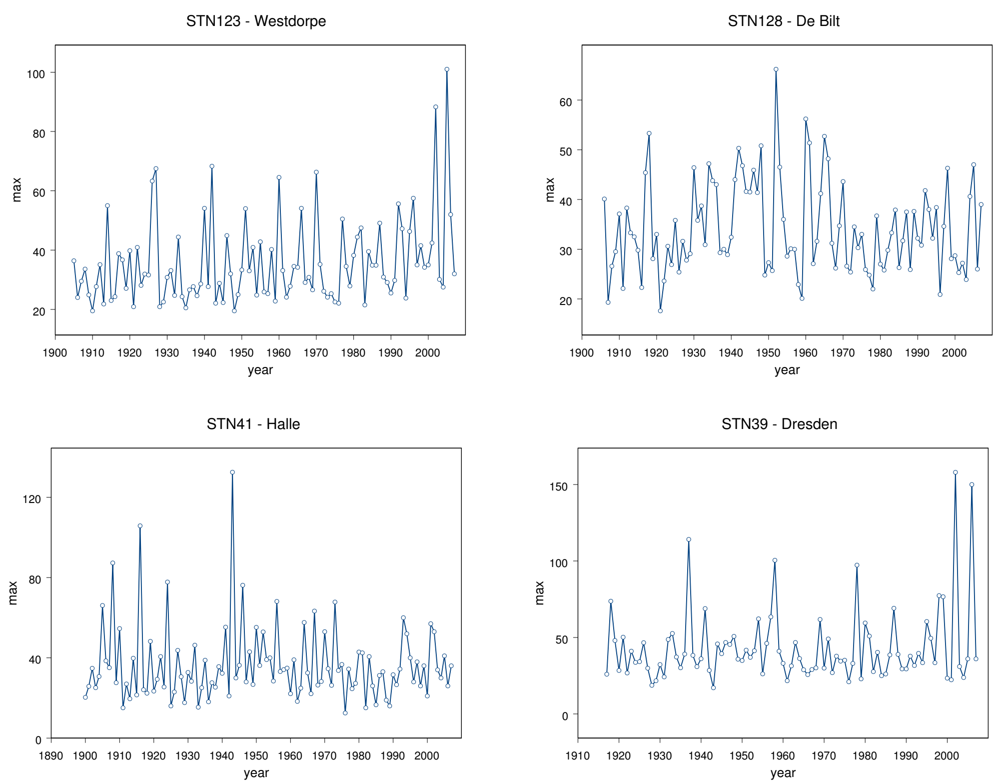
Figure 6 displays yearly maxima plots for several stations on the basis of available data. We get a mixed picture. In *STN41-Halle*, for instance, precipitation does not seem to be as severe now as in the first half of the 20th century anymore, whereas *STN39-Dresden* shows increasingly annual maxima with the largest peak of $158\,mm$ of rain, spot on the catastrophic event of 12 August 2002.

As long as there is at least one day in between, there is not much dependence in the amount of rainfall on two different days. For each station we select first the highest observation. Then we remove the observations on the day before and after. Next we select the highest observation from the remaining data, etc. This goes on until we have selected 70 days or the threshold of $1\,mm$ is reached. That way we get a sequence of higher order statistics from i.i.d. data. Table 1 displays the number of rain days (i.e. with at least $1\,mm$ of rain) per station.

Table 1: Total number of rain days in the period from 1918 to 2007 for each selected station.

STN	Name	Country	Lat.	Lon.	Rain days
36	Bamberg	D	48°49'N	11°33'E	6276
37	Berlin	D	52°31'N	13°20'E	6270
39	Dresden	D	51°31'N	13°44'E	6293
40	Frankfurt	D	50°6'N	08°40'E	6069
41	Halle	D	51°28'N	11°57'E	6230
42	Hamburg	D	53°33'N	09°59'E	6125
44	Hohenpeissenberg	D	47°48'N	10°59'E	6230
48	München	D	48°08'N	11°34'E	6153
49	Münster	D	52°59'N	07°41'E	5904
52	Stuttgart	D	48°46'N	09°46'E	6289
119	Ter Apel	NL	52°53'N	07°04'E	6069
120	Heerde	NL	52°24'N	06°03'E	6300
121	Winterswijk	NL	51°59'N	06°42'E	6298
122	Kerkwerve	NL	51°40'N	03°51'E	6276
123	Westdorpe	NL	51°13'N	03°52'E	6300
125	Roermond	NL	51°11'N	05°58'E	6300
128	De Bilt	NL	51°06'N	05°11'E	6299
129	Eelde	NL	53°08'N	06°35'E	6300

Figure 6: Some yearly maxima of daily rainfall.



Since we are not looking for a spatial trend now we shall make a study of the highest daily rainfall amounts in the 90 year period for each station separately. At each location, for $\hat{\gamma}^+$ (in connection with $\hat{c}^{(1)}$) we use Hill's estimator and for $\hat{\gamma}$ and $\hat{a}_0$ (for $\hat{c}^{(2)}$) we use the moment estimator (cf. sections 3.5 and 4.2 of de Haan and Ferreira (2006))

The point estimation of the extreme value index γ and trend estimation is conducted with the same number of upper order statistics k, just as prescribed in each definition of $\hat{c}^{(r)}$, $r = 1, 2, 3$, introduced in (5), (6), and (7), respectively.

In order to have enough tail related rain measurements per time point we found reasonable to take consecutive intervals of 5 years over the 90-year span. The disjoint intervals serve as our time points indexed by $j = 0, 1, 2, \ldots, 17 = m$.

For the purpose of data analysis, the gauging stations have been divided into two groups, determined by their alignments in the general climate characteristics (according to the Köppen-Geiger climate classification system, see e.g. Kottek et al. (2006)). Each selected station in Germany was classified as either *humid oceanic* or *humid continental*. All stations across The Netherlands are classified as *humid oceanic*. Although we are not looking for spatial coherence we hope to benefit from this information to get a more systematic presentation of our results.

Estimates of γ in a vicinity of 0.1 often emerge in connection with the extremal behavior of distributions underlying rainfall records (see e.g. Buishand et al. (2008), p.239; also Mannshardt-Shamseldin et al. (2010), p.492). This seems to hold for most of the considered stations although there is a lot of variation.

Figure 7 includes sample paths of the three proposals for estimating the tail trend parameter $c \in \mathbb{R}$ for some typical gauging stations. As already discussed, we shall handle estimation of c by screening plots as in Figure 7 for plateaus of stability in the early part of the graphs pertaining to the smoother estimator $\hat{c}^{(2)}$, coherent with the path patterns of $\hat{c}^{(1)}$ and $\hat{c}^{(3)}$ whenever possible.

Table 2 contains the estimated values of c for each station by their increasing order of magnitude. Standard errors are also provided. Bearing in mind the simulation results from section 3, here we shall confine attention to the leading estimator $\hat{c}^{(2)}$. Since the combined moment estimator $\hat{\gamma} = \hat{\gamma}_{n,k}$ is a consistent estimator for γ, the asymptotic standard error of $\hat{c}^{(2)} = \hat{c}$ can be estimated by

$$\begin{aligned} \widehat{\text{s.e.}}(\hat{c}^{(2)}) \quad := \quad & \frac{1}{\sqrt{k}} \Big\{ \sum_{j=1}^{m} S_j^2 \Big[\frac{(1 - e^{-\hat{\gamma}\hat{c}s_j} - \hat{\gamma}\hat{c}s_j)^2}{\hat{\gamma}^4} \, \frac{\sigma_\Gamma^2(\hat{\gamma})}{m} \\ & \qquad + 1 + e^{-2\hat{\gamma}\hat{c}s_j} + \frac{(1 - e^{-\hat{\gamma}\hat{c}s_j})^2}{\hat{\gamma}^2} \, \sigma_{A_0}^2(\hat{\gamma}) \Big] \Big\}^{\frac{1}{2}}, \end{aligned}$$

where $S_j := s_j / \sum_{j=1}^{m} s_j^2$,

$$\sigma_\Gamma^2(\gamma) := \begin{cases} 1 + \gamma^2, & \gamma \geq 0, \\ \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)}, & \gamma < 0 \end{cases}$$

Figure 7: Sample path of the overall moment estimator for γ and sample trajectories of $\hat{c}^{(r)}$, $r = 1, 2, 3$, as a function of the same number k of top observations for each 5-year interval between 1918 and 2007 for several stations.

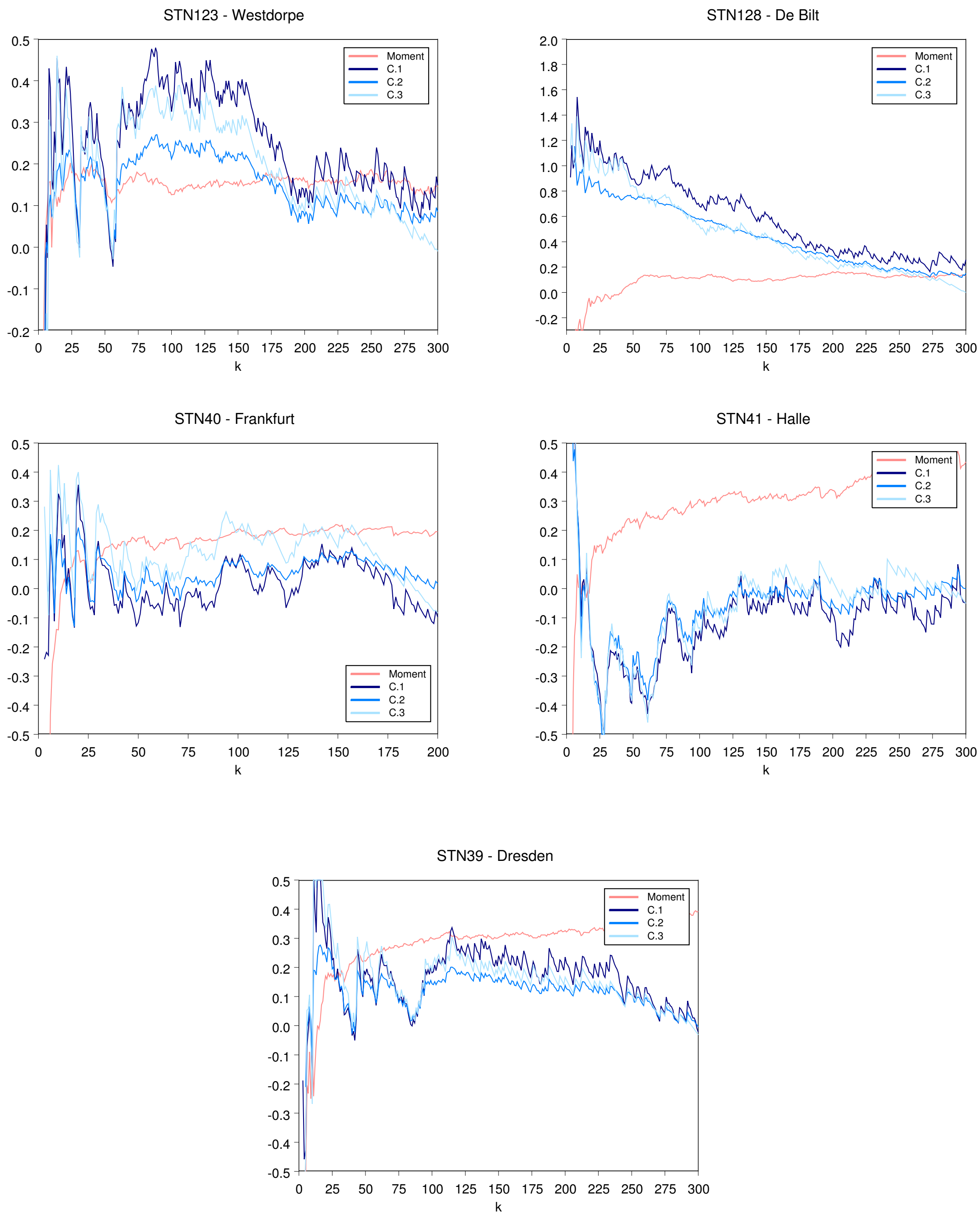


Table 2: Station-wise estimates for the trend parameter c by adopting estimator $\hat{c}^{(2)}$.

Continental Stations

Station	$\hat{c}$	$\widehat{\text{s.e.}}(\hat{c})$	$\hat{\gamma}$
STN 41	-0.2	0.095	0.2
STN 44	-0.1	0.101	0.15
STN 48	-0.1	0.109	0.13
STN 37	-0.06	0.122	0.2
STN 39	0.25	0.116	0.19
STN 36	0.32	0.116	0.18

Oceanic Stations

Station	$\hat{c}$	$\widehat{\text{s.e.}}(\hat{c})$	$\hat{\gamma}$
STN 49	0.16	0.122	0.12
STN 40	0.2	0.131	0.1
STN 123	0.2	0.113	0.19
STN 52	0.25	0.120	0.08
STN 122	0.55	0.131	0.1
STN 119	0.7	0.122	0.15
STN 120	0.7	0.132	0.15
STN 125	0.8	0.207	0.08
STN 128	0.8	0.261	-0.05
STN 121	0.9	0.290	0.05
STN 42	0.9	0.325	0.05
STN 129	1.0	0.221	0.07

and

$$\sigma^2_{A_0}(\gamma) := \begin{cases} 2+\gamma^2, & \gamma \geq 0, \\ \frac{2-16\gamma+51\gamma^2-69\gamma^3+50\gamma^4-24\gamma^5}{(1-2\gamma)(1-3\gamma)(1-4\gamma)}, & \gamma < 0. \end{cases}$$

It remains to assess whether the stations with near zero estimates in fact have a null trend. Examples are *STN37–Berlin*, *STN48–München*, *STN49–Münster*, *STN52–Stuttgart* and *STN39–Dresden*. The last site we refer to is *STN40–Frankfurt*, where testing for the presence of a trend is also of practical importance given the poor circumstances involving the estimation of the parameter c: the erratic sample paths displayed by the three estimators often cross the $c = 0$ line (cf. Figure 7). In the case of *STN40–Frankfurt* it seems difficult to find a “plateau of stability” in Figure 7; the estimate $\hat{c} = 0.2$ is rather uncertain. Therefore, we aim at a more definite decision on the value of c by means of a testing procedure upon *STN40–Frankfurt* in particular.

In order to tackle the problem of testing the presence of a trend in time, i.e., the problem of testing hypothesis

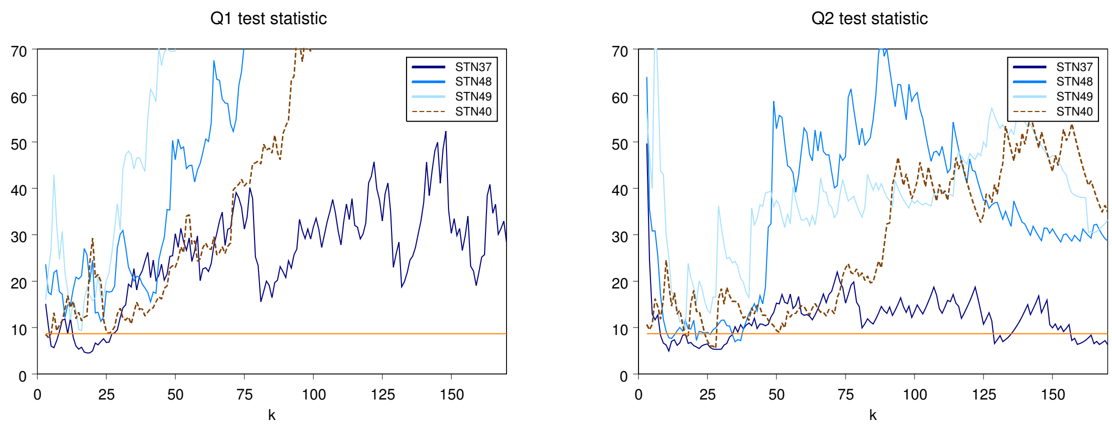
$$H_0 : c = 0 \quad versus \quad H_1 : c \neq 0, \tag{17}$$

we shall use $Q^{(r)}_{m,n}$ from corollary 2 as our test statistics. Hence, for $r = 1, 2$, the null hypothesis $H_0 : c = 0$ is rejected in favor of the two-sided alternative $H_1 : c \neq 0$ if $Q^{(r)}_{m,n} > q_{m,1-\alpha}$, where $q_{m,1-\alpha}$ denotes the $(1-\alpha)$-quantile of the chi-square distribution with m degrees of freedom.

Figure 8 depicts the sample trajectories pertaining to the two-sided test statistics $Q^{(1)}_{m,n}$ and $Q^{(2)}_{m,n}$ in companion with critical values at the nominal size $\alpha = 0.05$ and with respect to the referenced stations *STN37–Berlin*, *STN48–München*, *STN49–Münster* and *STN40–Frankfurt*. It seems that the hypothesis of no trend has to be rejected in case of *STN40–Frankfurt*. The two tests also ascertain a non-null trend for *STN48–München* and *STN49–Münster*. There is no evidence of a particular trend at *STN37–Berlin*.

Broadly speaking, the stations in the Oceanic group present higher positive values of c (but not always smaller values of γ) than the Continental stations. Overall, fairly positive values of c may be interpreted as

Figure 8: Sample paths of $Q^{(1)}_{18,n}$ and $Q^{(2)}_{18,n}$ test statistics and corresponding critical values for the two-sided test at a significance level $\alpha = 0.05$ ($\chi^{-1}_{0.95}(17) = 8.67$) for several stations across Germany.



being influenced by the ocean.

Allen and Ingram (2002) describe how the intensity of extreme rainfall events depends on the availability of moisture. Because moisture availability is constraint on temperature (through the Clausius-Clapeyron relationship), an increase in rainfall extremes is expected in a warming climate. Lenderink et al. (2009) show that higher increases can be expected at locations that are under the influence of the sea. Increasing sea surface temperatures contribute to higher rainfall amounts. The results obtained in this study for the behavior of extreme rainfall at locations in Germany and The Netherlands are consistent with these findings. Overall, the Oceanic group of stations shows a stronger increase in extreme rainfall than the Continental group of stations.

Next we discuss the interpretation of c. If $c = 1.0$ then in view of the fact that s is measured in periods of 5 years, the probability of really heavy rainfall increases approximately by 12.5% during each decade. Similarly if $c = 0.1$ then the probability of heavy rainfall increases in the same period approximately by 2%. These results are in good agreement with the positive trend of 3% per decade found by Zolina et al. (2009) for the second half of the century (1950-2000) using a different metric.

5 Proofs

We shall use the following representation:

$$\left\{X_{n-i,n}(s_j)\right\}_{i=1}^{n} \stackrel{d}{=} \left\{U_{s_j}(Y_{n-i,n}(s_j))\right\}_{i=1}^{n}$$

where $\{Y_{n-i,n}(s_j)\}_{i=1}^{n}$ are the $n-$th order statistics from the distribution function $1-x^{-1}$, $x \geq 1$, independently for $j = 1, 2, \ldots, m$.

Proof of consistency

For the consistency of $\hat{c}^{(1)}$ note that $(k/n)\, Y_{n-k,n}(s_j) \xrightarrow{P} 1$, $n \to \infty$ (cf. de Haan and Ferreira (2006), Corollary 2.2.2) and that $\lim_{t\to\infty} U_s(tx)/U_s(t) = x^{\gamma^+}$ locally uniformly for $x > 0$. Hence for $j = 1,\, 2, \ldots, m$

$$\log X_{n-k,n}(s_j) - \log U_{s_j}\big(\frac{n}{k}\big) \underset{n\to\infty}{\xrightarrow{P}} 1.$$

The rest is easy.

Similarly with respect to $\hat{c}^{(2)}$ we get that

$$\frac{X_{n-k,n}(s_j) - U_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)} \underset{n\to\infty}{\xrightarrow{P}} 0. \tag{18}$$

This limit relation combined with relation (3) leads directly to the consistency of $\hat{c}^{(2)}$.

With respect to the consistency of $\hat{c}^{(3)}$, we begin by noting that the domain of attraction condition

$$\lim_{n\to\infty} \frac{n}{k}\, P\Big\{X_l(0) > U_0\big(\frac{n}{k}\big) + x\, a_0\big(\frac{n}{k}\big)\Big\} = (1+\gamma x)^{-1/\gamma}$$

for all $l = 1, 2, \ldots$, combined with (1) implies

$$\lim_{n\to\infty} e^{-cs_j} \frac{n}{k}\, P\Big\{X_l(s_j) > U_0\big(\frac{n}{k}\big) + x\, a_0\big(\frac{n}{k}\big)\Big\} = (1+\gamma x)^{-1/\gamma},$$

$j = 1, 2, \ldots, m$. Hence the characteristic functions converge:

$$\begin{aligned}
& E \exp\Big\{ i \frac{t}{k} \sum_{l=1}^{n} I_{\{X_l(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \Big\} \\
= \quad & \Big(E \exp\Big\{ i \frac{t}{k}\, I_{\{X_1(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \Big\} \Big)^n \\
= \quad & \Big(e^{i\frac{t}{k}}\, P\Big\{ X_1(s_j) > U_0\big(\frac{n}{k}\big) + x\, a_0\big(\frac{n}{k}\big) \Big\} \\
& \qquad + 1 - P\Big\{ X_1(s_j) > U_0\big(\frac{n}{k}\big) + x\, a_0\big(\frac{n}{k}\big) \Big\} \Big)^n \\
= \quad & \Big(1 + e^{cs_j} k \big(e^{it/k} - 1 \big)\, \frac{e^{-cs_j}(n/k)\big[1 - F_{s_j}(U_0(n/k) + x a_0(n/k)) \big]}{n} \Big)^n \\
\underset{n\to\infty}{\longrightarrow} \quad & \exp\Big\{ it\, e^{cs_j} (1 + \gamma x)^{-1/\gamma} \Big\},
\end{aligned}$$

for every $t \in \mathbb{R}$. Owing to Lévy's continuity theorem, the latter implies

$$\frac{1}{k} \sum_{i=1}^{n} I_{\{X_i(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \xrightarrow[n\to\infty]{P} e^{cs_j} (1 + \gamma\, x)^{-1/\gamma}.$$

Next use (18). □

For the proof of the asymptotic normality we need an auxiliary result.

Lemma 3 *Assume conditions* (13) *and* (14). *Define*

$$\begin{aligned}
\alpha_s(t) &:= e^{cs\rho} \alpha(t), \\
a_s(t) &:= e^{cs\gamma} a_0(t) \Big(1 + \alpha(t)\, \frac{e^{cs\rho} - 1}{\rho} \Big).
\end{aligned} \tag{19}$$

Then for $s \in \mathbb{R}$ *and* $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_s(tx) - U_s(t)}{a_s(t)} - \frac{x^\gamma - 1}{\gamma}}{\alpha_s(t)} = H_{\gamma,\rho}(x). \tag{20}$$

Proof: For simplicity we write d for e^{cs} in this proof.

First note that (13) implies that

$$\begin{aligned}
\frac{\frac{U_s(tx) - U_s(t)}{a_0(t)} - d^\gamma \frac{x^\gamma - 1}{\gamma}}{\alpha(t)} \quad = \quad & \frac{\frac{U_0(tx) - U_0(t)}{a_0(t)} - \frac{x^\gamma - 1}{\gamma}}{\alpha(t)} + \frac{\frac{a_0(tx)}{a_0(t)} - x^\gamma}{\alpha(t)}\, \frac{d^\gamma - 1}{\gamma} \\
& + \Big(\frac{\alpha(tx)\, a_0(tx)}{\alpha(t)\, a_0(t)} - 1 \Big) H_{\gamma,\rho}(d) \big(1 + o(1) \big)
\end{aligned}$$

converges to

$$H_{\gamma,\rho}(x) + x^{\gamma}\,\frac{x^{\rho}-1}{\rho}\,\frac{d^{\gamma}-1}{\gamma} + (x^{\gamma+\rho}-1)H_{\gamma,\rho}(d).$$

Next write

$$\frac{\frac{U_s(tx)-U_s(t)}{a_s(t)} - \frac{x^{\gamma}-1}{\gamma}}{\alpha_s(t)} = \frac{\alpha(t)}{\alpha_s(t)}\left(\frac{\frac{U_s(tx)-U_s(t)}{a_0(t)} - d^{\gamma}\frac{x^{\gamma}-1}{\gamma}}{\alpha(t)}\,\frac{a_0(t)}{a_s(t)} + \frac{x^{\gamma}-1}{\gamma}\,\frac{\frac{d^{\gamma}a_0(t)}{a_s(t)}-1}{\alpha(t)}\right).$$

This converges to

$$d^{-(\gamma+\rho)}\Big(H_{\gamma,\rho}(x) + x^{\gamma}\,\frac{x^{\rho}-1}{\rho}\,\frac{d^{\gamma}-1}{\gamma} + (x^{\gamma+\rho}-1)H_{\gamma,\rho}(d)\Big) - \frac{x^{\gamma}-1}{\gamma}\,\frac{1-d^{-\rho}}{\rho}\Big)$$

which is equal to $H_{\gamma,\rho}(x)$. ❏

Remark 4 *The analogue of Lemma 3 stemming from conditions* (9) *and* (10) *– i.e.* $\gamma > 0$ *– holds with the auxiliary function* $\beta_s(t) := e^{cs\widetilde{\rho}}\beta(t)$*. This leads to the following relation for every* $s \in \mathbb{R}$*,*

$$\lim_{t\to\infty}\frac{\frac{U_s(tx)}{U_s(t)} - e^{c\gamma^{+}s}}{\beta_s(t)} = x^{\gamma^{+}}\,\frac{x^{\widetilde{\rho}}-1}{\widetilde{\rho}}, \quad x>0.$$

Proof of asymptotic normality

We write

$$\begin{aligned}
\sqrt{k}\,\big(\hat{c}^{(1)}-c\big) \;=\; & \sqrt{k}\,\Big(\frac{1}{\hat{\gamma}^{+}_{n,k}} - \frac{1}{\gamma^{+}}\Big)\,\hat{\gamma}^{+}_{n,k}\,\hat{c}^{(1)} + \Big(\gamma^{+}\sum_{j=1}^{m}s_j^2\Big)^{-1} \\
& \times\Big\{\sum_{j=1}^{m}s_j\Big[\sqrt{k}\,\Big(\log X_{n-k,n}(s_j) - \log U_{s_j}\big(\frac{n}{k}\big)\Big) \\
& \quad - \sqrt{k}\,\Big(\log X_{n-k,n}(0) - \log U_0\big(\frac{n}{k}\big)\Big) \\
& \quad\quad + \sqrt{k}\,\beta\big(\frac{n}{k}\big)\,\frac{\log U_{s_j}(n/k) - \log U_0(n/k)}{\beta(n/k)}\Big]\Big\}.
\end{aligned}$$

The result follows from (11) and (10).

For $\hat{c}^{(2)}$ it is sufficient to consider

$$\sqrt{k}\,\Big\{\log\Big(1+\hat{\gamma}_{n,k}\,\frac{X_{n-k,n}(s_j)-X_{n-k,n}(0)}{\hat{a}_0\big(\frac{n}{k}\big)}\Big)^{\frac{1}{\hat{\gamma}_{n,k}}} - cs_j\Big\}$$

for $j = 1, 2, \ldots, m$ where $\hat{\gamma}_{n,k} = 1/m \sum_{j=1}^{m} \hat{\gamma}_{n,k}(s_j)$. We use Cramér's delta method.

$$\frac{\partial}{\partial \gamma} \log(1+\gamma x)^{\frac{1}{\gamma}} = \frac{1}{\gamma^2}\Big(\frac{\gamma x}{1+\gamma x} - \log(1+\gamma x)\Big)$$

(which is $x^2/2$ for $\gamma = 0$) and

$$\frac{\partial}{\partial x} \log(1+\gamma x)^{\frac{1}{\gamma}} = \frac{1}{1+\gamma x}$$

($= 1$ for $x = 0$). Further we write

$$\begin{aligned}
&\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)} \\
= \quad & \frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)} \left\{ \frac{a_{s_j}\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} \frac{X_{n-k,n}(s_j) - U_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)} \right. \\
& \qquad \left. - \frac{X_{n-k,n}(0) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} + \frac{U_{s_j}\left(\frac{n}{k}\right) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} \right\},
\end{aligned}$$

implying

$$\begin{aligned}
&\sqrt{k}\Big(\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma} \Big) \\
= \quad & \frac{e^{c\gamma s_j} - 1}{\gamma} \sqrt{k} \left(\frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)} - 1 \right) + \frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)} \left\{ \frac{a_{s_j}\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} \sqrt{k}\, \frac{X_{n-k,n}(s_j) - U_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)} \right. \\
& \qquad \left. - \sqrt{k}\, \frac{X_{n-k,n}(0) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} + \sqrt{k} \left(\frac{U_{s_j}\left(\frac{n}{k}\right) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma} \right) \right\}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sqrt{k}\Big(\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma} \Big) \\
\xrightarrow[n\to\infty]{d} \quad & -\frac{e^{c\gamma s_j} - 1}{\gamma} A(0) + e^{c\gamma s_j} B(s_j) - B(0) + \lambda\, H_{\gamma,\rho}(e^{c s_j}).
\end{aligned}$$

Next we apply the delta method:

$$
\begin{aligned}
&\sqrt{k}\Big\{\log\Big(1+\hat\gamma_{n,k}\frac{X_{n-k,n}(s_j)-X_{n-k,n}(0)}{\hat a_0\big(\frac{n}{k}\big)}\Big)^{\frac{1}{\hat\gamma_{n,k}}} \\
&\qquad\qquad\qquad - \log\Big(1+\gamma\,\frac{e^{c\gamma s_j}-1}{\gamma}\Big)^{\frac{1}{\gamma}}\Big\} \\
\xrightarrow[n\to\infty]{d}\quad & \Big[\frac{\partial}{\partial\gamma}\,\log(1+\gamma x)^{\frac{1}{\gamma}}\Big]_{x=\frac{e^{c\gamma s_j}-1}{\gamma}}\cdot m^{-1}\sum_{j=1}^m \Gamma(s_j) \\
&\quad + \Big[\frac{\partial}{\partial x}\,\log(1+\gamma x)^{\frac{1}{\gamma}}\Big]_{x=\frac{e^{c\gamma s_j}-1}{\gamma}}\Big\{-\frac{e^{c\gamma s_j}-1}{\gamma}\,A(0) \\
&\qquad\qquad + e^{c\gamma s_j}B(s_j)-B(0)+\lambda\,H_{\gamma,\rho}(e^{cs_j})\Big\} \\
=\quad & \frac{1-e^{-c\gamma s_j}-c\gamma s_j}{\gamma^2}\,\frac{1}{m}\sum_{j=1}^m\Gamma(s_j)+e^{-c\gamma s_j}\Big\{e^{c\gamma s_j}B(s_j) \\
&\qquad\qquad - B(0)-\frac{e^{c\gamma s_j}-1}{\gamma}A(0)+\lambda\,H_{\gamma,\rho}(e^{cs_j})\Big\}.
\end{aligned}
$$

The result follows. □

Finally for $\hat c^{(3)}$ consider

$$
\frac{1}{k}\sum_{i=1}^n I_{\big\{X_i(s)>X_{n-k,n}(0)\big\}}=\frac{1}{k}\sum_{i=1}^n I_{\big\{\frac{X_i(s)-U_s^\star(n/k)}{a_s^\star(n/k)}>\frac{X_{n-k,n}(0)-U_s^\star(n/k)}{a_s^\star(n/k)}\big\}}. \tag{21}
$$

We write

$$
x_n(s):=\frac{X_{n-k,n}(0)-U_s^\star\big(\frac{n}{k}\big)}{a_s^\star\big(\frac{n}{k}\big)},
$$

with $U^\star$ and $a^\star$ from Corollary 2.3.7 of de Haan and Ferreira (2006), coupled with lemma 3,

$$
\begin{aligned}
U_s^\star(t) &:= \begin{cases} U_s(t)-\frac{e^{cs(\gamma+\rho)}}{\rho(\gamma+\rho)}\,a_0(t)\alpha(t), & \gamma+\rho\neq 0,\ \rho<0, \\ U_s(t), & \text{otherwise,}\end{cases} \\
a_s^\star(t) &:= \begin{cases} e^{cs\gamma}a_0(t)\big(1-\frac{1}{\rho}\alpha(t)\big), & \rho<0, \\ e^{cs\gamma}a_0(t)\big[1-\big(\frac{e^{cs\rho}}{\gamma\rho}+cs\big)\alpha(t)\big], & \rho=0\neq\gamma, \\ a_0(t)\big(1+cs\,\alpha(t)\big), & \gamma=\rho=0,\end{cases}
\end{aligned}
$$

and write (21) as

$$
\frac{n}{k}\,\Big\{1-F_n^{(s)}\Big(U_s^\star\big(\frac{n}{k}\big)+x_n(s)\,a_s^\star\big(\frac{n}{k}\big)\Big)\Big\}.
$$

where $F_n^{(s)}$ is the empirical distribution function of the random sample $X_1(s),X_2(s),\dots,X_n(s)$. We consider

the two parts separately. By (14) and Theorem 2.4.2 of de Haan and Ferreira (2006),

$$\begin{aligned}
&\sqrt{k}\Big(x_n(s) - \frac{e^{-cs\gamma}-1}{\gamma}\Big) \\
=\;& \frac{a_0\big(\frac{n}{k}\big)}{a_s^\star\big(\frac{n}{k}\big)}\left\{\sqrt{k}\,\frac{X_{n-k,n}(0)-U_0\big(\frac{n}{k}\big)}{a_0\big(\frac{n}{k}\big)} - \sqrt{k}\left(\frac{U_s\big(\frac{n}{k}\big)-U_0\big(\frac{n}{k}\big)}{a_0\big(\frac{n}{k}\big)} - \frac{e^{cs\gamma}-1}{\gamma}\right)\right\} \\
&\qquad - \sqrt{k}\left(\frac{a_0\big(\frac{n}{k}\big)}{a_s^\star\big(\frac{n}{k}\big)} - e^{-cs\gamma}\right)\frac{e^{cs\gamma}-1}{\gamma} - \sqrt{k}\,\frac{U_s^\star\big(\frac{n}{k}\big)-U_s\big(\frac{n}{k}\big)}{a_s^\star\big(\frac{n}{k}\big)} \\
&\xrightarrow[n\to\infty]{d} e^{-cs\gamma}\Big\{W_n^{(0)}(1) - \lambda H_{\gamma,\rho}(e^{cs}) + \lambda\, b^\star(s)\Big\}
\end{aligned} \tag{22}$$

where

$$b^\star(s) = \begin{cases} \frac{1}{\rho}\big(\frac{e^{cs(\gamma+\rho)}}{\gamma+\rho} - \frac{e^{cs\gamma}-1}{\gamma}\big), & \gamma+\rho\neq 0,\ \rho<0, \\ -\frac{1}{\rho}\frac{e^{cs\gamma}-1}{\gamma}, & \rho=-\gamma<0, \\ \frac{e^{cs\gamma}-1}{\gamma}\big(cs-\frac{1}{\gamma}\big), & \rho=0\neq\gamma, \\ (cs)^2, & \gamma=\rho=0. \end{cases}$$

Further by Theorem 5.1.2 of de Haan and Ferreira (2006), since $x = x_n$ is asymptotically constant,

$$\begin{aligned}
&\sqrt{k}\left\{\frac{n}{k}\left(1 - F_n^{(s)}\Big(U_s^\star\big(\frac{n}{k}\big) + x_n(s)\,a_s^\star\big(\frac{n}{k}\big)\Big)\right) - \big(1+\gamma\,x_n(s)\big)^{-\frac{1}{\gamma}}\right\} \\
&\quad - W_n^{(s)}\Big(\big(1+\gamma\,x_n(s)\big)^{-\frac{1}{\gamma}}\Big) \\
&\qquad - \sqrt{k}\,\alpha_s^\star\big(\frac{n}{k}\big)\big(1+\gamma\,x_n(s)\big)^{-\frac{1}{\gamma}-1}\Psi_{\gamma,\rho}\Big(\big(1+\gamma\,x_n(s)\big)^{\frac{1}{\gamma}}\Big) \xrightarrow[n\to\infty]{P} 0
\end{aligned}$$

for a sequence of standard Brownian motions $\{W_n^{(s)}(t)\}_{t\geq 0}$ and with

$$\alpha_s^\star(t) = \begin{cases} \frac{e^{cs\rho}}{\rho}\,\alpha(t), & \rho<0, \\ e^{cs\rho}\alpha(t), & \rho=0, \end{cases}$$

and

$$\Psi_{\gamma,\rho}(x) := \begin{cases} \frac{x^{\gamma+\rho}}{\gamma+\rho}, & \gamma+\rho\neq 0,\ \rho<0, \\ \log x, & \gamma+\rho=0,\ \rho<0, \\ \frac{1}{\gamma}\,x^\gamma\log x, & \rho=0\neq\gamma, \\ \frac{1}{2}\,(\log x)^2, & \rho=0=\gamma. \end{cases}$$

Since by (22)

$$\begin{aligned}
&\sqrt{k}\left\{\big(1+\gamma\,x_n(s)\big)^{-\frac{1}{\gamma}} - \Big(1+\gamma\,\frac{e^{-cs\gamma}-1}{\gamma}\Big)^{-\frac{1}{\gamma}}\right\} \\
\xrightarrow[n\to\infty]{d}\;& -e^{cs(\gamma+1)}\Big(e^{-cs\gamma}\big\{W^{(0)}(1) - \lambda H_{\gamma,\rho}(e^{cs}) + \lambda\, b^\star(s)\big\}\Big),
\end{aligned}$$

$$\sqrt{k}\left\{\frac{n}{k}\left(1-F_n^{(s)}\Big(U_s^\star\big(\frac{n}{k}\big)+x_n(s)\,a_s^\star\big(\frac{n}{k}\big)\Big)\right)-e^{cs}\right\}$$
$$\xrightarrow[n\to\infty]{d}\quad W^{(s)}(e^{cs})-e^{cs}W^{(0)}(1)+\lambda e^{cs}\Big(H_{\gamma,\rho}(e^{cs})+e^{cs\gamma}\Psi_{\gamma,\rho}(e^{-cs})-b^\star(s)\Big).$$

Proof: [of Corollary 2] By virtue of Rao's theorem (Rao (1973), Section 3.b.4; see also Serfling (2002), p.128) pertaining to quadratic forms of asymptotically normal random vectors, statements (15) and (16) follow immediately from the theorem. ❏

A Sketch of alternative approaches

The subject of extreme value theory (EVT) is the study of the right (or left) tail of a probability distribution near the endpoint. Hence by nature EVT is an asymptotic theory. The basic assumption is

$$\lim_{n\to\infty} P\Big\{\frac{\max(X_1,\,X_2,\dots,X_n)-b_n}{a_n}\le x\Big\}=\exp\big\{-(1+\gamma\,x)^{-1/\gamma}\big\}, \tag{23}$$

where $X_1, X_2, \dots$ are i.i.d. random variables. It follows that the limit distribution has only one parameter, the shift $b_n \in \mathbb{R}$ and scale $a_n > 0$ are not parameters of the limit distribution. They depend essentially on the distribution of X_1.

When it comes to statistics there are three basic methods:

1. Yearly maxima (or block maxima). Over a number of years one takes the yearly maximum. Since the yearly maximum is taken over many underlying random variables (albeit not i.i.d.) the assumption is that the yearly maximum M_j can be considered the maximum over a large number n of i.i.d random variables so that

$$P\big\{M_j\le x\big\}\approx\exp\Big\{-\Big(1+\gamma\,\frac{x-b_n}{a_n}\Big)^{-\frac{1}{\gamma}}\Big\}$$

where n is unknown. The random variables M_j are i.i.d.. The right hand-side can then be interpreted as a parametric model (GEV: Generalized Extreme Value distribution) so that e.g. the method of maximum likelihood can be applied.

The interpretation of b_n is: the level that has a return period (the mean time between consecutive exceedances of the level) of $e/(e-1)\approx 1.58$ years. Hence there is no direct intuitive meaning for b_n. Also the behavior of b_n as $n\to\infty$ can not be found. This method carries a bias stemming from replacing an approximate equality with a firm equality. In contrast to the next case it seems difficult to control that bias.

2. Peaks over threshold. The basic assumption (23) implies that with $b(t)=b_{[t]}$, $a(t)=a_{[t]}$ and $[t]$ the

integer part of t for $x > 0$

$$\lim_{t\to\infty} P\Big\{\frac{X_1 - b(t)}{a(t)} > x | X_1 > b(t)\Big\} = (1+\gamma x)^{-1/\gamma}. \tag{24}$$

Select out of n i.i.d. observations the ones that are larger than $b(t)$. These are approximately i.i.d. and (when normalized) follow approximately the GPD distribution $1-(1+\gamma\, x)^{-1/\gamma}$ (Pickands (1975)).

One can take for $b(t)$ one of the order statistics $X_{n-k,n}$. In order to get meaningful results we need to have $k = k_n$ and $k_n \to \infty$, $k(n)/n \to 0$ as $n \to \infty$. Then $X_{n-k,n}$ is close to $b(n/k)$, i.e., the quantile $F^{\leftarrow}(1-k/n)$.

Again, since for $x > b(t)$

$$P\big\{X_1 > x | X_1 > b(t)\big\} \approx \Big(1 + \frac{x-b(t)}{a(t)}\Big)^{-1/\gamma},$$

one can consider the right hand-side as a parametric model so that e.g. the method of maximum likelihood can be applied.

Next one can prove that the obtained estimators are consistent and asymptotically normal as n, the number of observations, tends to infinity. That is, the vector

$$\sqrt{k}\left(\frac{\hat{a}\big(\frac{n}{k}\big)}{a\big(\frac{n}{k}\big)} - 1,\, \hat{\gamma}_{n,k} - \gamma\right)$$

has asymptotically a normal distribution(Smith (1987); Drees et al. (2004); Zhou (2009)). There are also methods to minimize the bias by choosing the number k appropriately.

3. Point process convergence. Suppose that the basic assumption holds. Take the point process on $\mathbb{R}^2$ with points

$$\Big\{\Big(\frac{i}{n}, \frac{X_i - b_n}{a_n}\Big)\Big\}_{i=1}^{n}. \tag{25}$$

This point process converges in distribution to a Poisson point process on $(0,1)\times\mathbb{R}$ with intensity measure $dt\cdot(1+\gamma x)^{-1/\gamma-1}dx$ (cf. Pickands (1971)). Note that the intensity measure is unbounded. Those points in (25) for which $(X_i - b_n)/a_n$ exceeds some threshold u are approximately points from a Poisson point process with (finite) parametric intensity measure so that the method of maximum likelihood can be applied supplying estimators for γ, b_n and a_n (see Smith (1989), cf. Coles (2001), Chapter 7). No asymptotic behavior ($n \to \infty$, $u = u_n$ decreasing) seems to be known for these estimators.

The three methods have been explained in detail in the book of Coles (2001). A trend in the EVT analysis has been considered in all three methods.

1$'$. Chapter 6 of Coles (2001) book treats trends in the block maxima / GEV setup. One considers time

points $j = 0, 1, 2, \ldots$ and assumes (for example)

$$b_n(j) = b_n(0) + j\,c$$

or/and

$$\log a_n(j) = \log a_n(0) + j\,c'.$$

As we saw before, b_n is the level that has a return period of just $e/(e-1)$ years. The scale a_n can be interpreted with some liberty as a derivative, i.e., speed of change of location. The interpretation of both seems less straightforward than that of (1).

There is also another complication. If one is interested in the location parameter over a longer period, say, of 2 years i.e. n replaced with $2n$, the relation is

$$\begin{aligned} b_{2n}(j) &\approx b_n(j) + a_n(j)\,\frac{2^\gamma - 1}{\gamma} \\ &\approx b_{2n}(0) + j\,c + (a_n(j) - a_n(0))\,\frac{2^\gamma - 1}{\gamma}. \end{aligned}$$

This is no longer a linear trend in general. Note that our framework, combining trends in location and scale, is not bound to a certain period.

2$'$.Peaks over threshold. Davison and Smith (1990) consider a linear trend in both γ and $a(n/k)$. Coles (2001), p.119, considers a linear trend in $\log a(n/k)$. Estimation is done by maximum likelihood. No asymptotic analysis of the quality of the estimators as the number of observations tends to infinity is made. Again the interpretation of a trend in $a(n/k)$ or $\log a(n/k)$ seems difficult.

Hall and Tajvidi (2000) consider a nonlinear trend in γ. The method is likelihood based. No large sample results are given.

3$'$. Smith (1989) (c.f. Coles (2001) p.133 sqq.) considers a trend in the location

$$b_n(j) = b_n(0) + j\,c$$

(simplified) in the point process model. Estimation is by maximum likelihood. No asymptotic analysis ($n \to \infty$) is made. Another possible problem is that the trend for $b_n(j)$ could get out of range if $\gamma < 0$.

In short: the present paper looks at changes in (tail) probabilities whereas in the literature changes in various quantiles have been considered. The two viewpoints are not equivalent.

References

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, 111:(D05109).

Allen, M. R. and Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419:224–232.

Buishand, T. A., de Haan, L., and Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Ann. Appl. Statist.*, 2:624–642.

Coles, S. (2001). *An Introduction to Statistical Modelling of Extremes*. Springer Verlag.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *J.R. Statist. Soc. B*, 52:393–442.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer.

Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.*, 14:1179–1201.

Groisman, P. Y., Knight, R. W., Easterling, D. R., Karl, T. R., Hegerl, G. C., and Razuvaev, V. N. (2005). Trends in intense precipitation in the climate record. *J. Climate*, 18:1326–1350.

Hall, P. and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.*, 15:153–167.

Hanel, M., Buishand, T. A., and Ferro, C. A. T. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *J. Geophys. Res.*, 114:(D15107).

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3:1163–1174.

Klein Tank, A. M. G. and Können, G. P. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *J. Climate*, 16:3665–3680.

Kottek, M., Grieser, J., Christoph, B., Rudolph, B., and Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.*, 15:259–263.

Lenderink, G., van Meijgaard, E., and Selten, F. (2009). Intense coastal rainfall in The Netherlands in response to high sea surface temperatures: analysis of the event of august 2006 from the perspective of a changing climate. *Clim. Dyn.*, 32:19–33.

Mannshardt-Shamseldin, E. C., Smith, R. L., Sain, S. R., Mearns, L. O., and Cooley, D. (2010). Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. *Ann. Appl. Statist.*, 4:484–502.

Pickands, J. (1971). The two-dimensional Poisson process and extremal processes. *J. Appl. Probab.*, 8:745–756.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York.

Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Statistics, Wiley, New York.

Smith, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.*, 15:1174–1207.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4:367–393.

Yee, T. W. and Stephenson, A. G. (2007). Vector generalized linear and additive extreme value models. *Extremes*, 10:119.

Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *J. Multivariate Anal.*, 100:794–815.

Zolina, O., Simmer, C., Belyaev, K., Kapala, A., and Gulev, S. (2009). Improving estimates of heavy and extreme precipitation using daily records from European rain gauges. *J. Hydrometeorol.*, 10:701–716.